I. AMENDMENTS

In the specification:

Please replace the paragraph beginning on page 3, line 12 with the following rewritten paragraph:

--Zinc finger proteins ("ZFPs") are proteins that can bind to DNA in a sequence-specific manner. Zinc fingers were first identified in the transcription factor TFIIIA from the oocytes of the African clawed toad, *Xenopus laevis*. ZFPs are widespread in eukaryotic cells. An exemplary motif characterizing one class of these proteins (C_2H_2 class) is -Cys-$(X)_{2-4}$-Cys-$(X)_{12}$-His-$(X)_{3-5}$-His (SEQ ID NO:1) (where X is any amino acid). A single finger domain is about 30 amino acids in length and several structural studies have demonstrated that it contains an alpha helix containing the two invariant histidine residues co-ordinated through zinc with the two cysteines of a single beta turn. To date, over 10,000 zinc finger sequences have been identified in several thousand known or putative transcription factors. ZFPs are involved not only in DNA-recognition, but also in RNA binding and protein-protein binding. Current estimates are that this class of molecules will constitute about 2% of all human genes.--

D2

Please replace the paragraph beginning on page 13, line 15 with the following rewritten paragraph:

–The term "zinc finger protein" or "ZFP" refers to a protein having DNA binding domains that are stabilized by zinc. The individual DNA binding domains are typically referred to as "fingers" A zinc finger protein has least one finger, typically two fingers, three fingers, four fingers, five fingers, or six fingers or more. Each finger binds from two to four base pairs of DNA, typically three or four base pairs of DNA. A zinc finger protein binds to a nucleic acid sequence called a target site or target segment. Each finger

D3

typically comprises an approximately 30 amino acid, zinc-coordinating, DNA-binding subdomain. An exemplary motif characterizing one class of these proteins (Cys_2His_2 class) is -Cys-$(X)_{2-4}$-Cys-$(X)_{12}$-His-$(X)_{3-5}$-His (SEQ ID NO:1) (where X is any amino acid). Studies have demonstrated that a single zinc finger of this class consists of an alpha helix containing the two invariant histidine residues co-ordinated with zinc along with the two cysteine residues of a single beta turn (*see, e.g.*, Berg & Shi, *Science* 271:1081-1085 (1996)).--

D3

On page 21, please replace the paragraph beginning on line 10 with the following:

--The ZFPs of the invention are engineered to recognize a selectable target site in the endogenous gene of choice. Typically, a backbone from any suitable C_2H_2 (SEQ ID NO:1) ZFP, such as SP-1, SP-1C, or ZIF268 is used as the scaffold for the engineered ZFP (*see, e.g.,* Jacobs *EMBO J.* 11:4507 (1992); Desjarlais & Berg, *PNAS* 90:2256-2260 (1993)). A number of methods can then be used to design and select a ZFP with high affinity for its target (e.g., preferably with a K_d of less than about 25 nM). As described above, a ZFP can be designed or selected to bind to any suitable target site in the target endogenous gene, with high affinity. Co-owned WO 00/42219, incorporated by reference herein in its entirety, comprehensively describes methods for design, construction, and expression of ZFPs for selected target sites.--

D4

On page 21, please replace the paragraph beginning on line 32 with the following:

--In a preferred embodiment, co-owned WO 00/42219 provides methods that select a target gene, and identify a target site within the gene containing one to six (or more) D-able sites (see definition below). Using these methods, a ZFP can then be synthesized that binds to the preselected site. These methods of target site selection are premised, in part, on the recognition that the presence of one or more D-able sites in a target segment confers the potential for higher binding affinity in a ZFP selected or

D5

D5 cont

designed to bind to that site relative to ZFPs that bind to target segments lacking D-able sites. Experimental evidence supporting this insight is provided in Examples 2-9 of co-owned WO 00/42219.--

On page 22, please replace the paragraph beginning on line 9 with the following:

--A D-able site or subsite is a region of a target site that allows an appropriately designed single zinc finger to bind to four bases rather than three of the target site. Such a zinc finger binds to a triplet of bases on one strand of a double-stranded target segment (target strand) and a fourth base on the other strand (*see* Figure 2 of co-owned WO 00/42219). Binding of a single zinc finger to a four base target segment imposes constraints both on the sequence of the target strand and on the amino acid sequence of the zinc finger. The target site with the target strand should include the "D-able" site motif 5' NNGK 3' (SEQ ID NO:41), in which N and K are convention IUPAC-IUB

D6

ambiguity codes. A zinc finger for binding to such a site should include an arginine residue at position -1 and an aspartic acid, (or less preferably a glutamic acid) at position +2. The arginine residues at position -1 interacts with the G residue in the D-able site. The aspartic acid (or glutamic acid) residue at position +2 of the zinc finger interacts with the opposite strand base complementary tot he K base in the D-able site. It is the interaction between aspartic acid (symbol D) and the opposite strand base (fourth base) that confers the name D-able site. AS is apparent from the D-able site formula, there are two subtypes of D-able sites; 5' NNGG 3' (SEQ ID NO:42) and 5' NNGT 3' (SEQ ID NO:43). For the former site, the aspartic acid or glutamic acid at position +2 of a zinc finger interacts with a C in the opposite strand to the D-able site. In the latter site, the aspartic acid or glutamic acid at position +2 of a zinc finger interacts with an A in the opposite strand of the D-able site. In general, NNGG (SEQ ID NO:42) is preferred over NNGT (SEQ ID NO:43).--

Please replace the paragraph beginning on page 23, line 12 with the following rewritten paragraph:

--In the formula 5'-NNx aNy bNzc-3', the triplets of NNx aNy and bNzc represent the triplets of bases on the target strand bound by the three fingers in a ZFP. If only one of x, y and z is a G, and this G is followed by a K, the target site includes a single D-able subsite. For example, if only x is G, and a is K, the site reads 5'-**NNG KN**y bNzc-3' with the D-able subsite highlighted. If both x and y but not z are G, and a and b are K, then the target site has two overlapping D-able subsites as follows: 5'-**NNG *KNG K***Nz c-3' (SEQ ID NO:2), with one such site being represented in bold and the other in italics. If all three of x, y and z are G and a, b, and c are K, then the target segment includes three D-able subsites, as follows 5'**NNG *KNG* <u>KNG K</u>**3' (SEQ ID NO:3), the D-able subsites being represented by bold, italics and underline.--

On page 27, please replace the paragraph beginning on line 8 with the following:

--The biochemical properties of the purified proteins, e.g., K_d, can be characterized by any suitable assay. In one embodiment, K_d is characterized via electrophoretic mobility shift assays ("EMSA") (Buratowski & Chodosh, *in Current Protocols in Molecular Biology* pp. 12.2.1.-12.2.7 (Ausubel ed., 1996); *see also* U.S. Patent No. 5,789,538; co-owned WO 00/42219 and Example 1, *infra*). Affinity is measured by titrating purified protein against a low amount of labeled double-stranded oligonucleotide target. The target comprises the nature binding site sequence (9 or 18 bp) flanked by the 3 bp found in the natural sequence. External to the binding site plus flanking sequence is a constant sequence. The annealed oligonucleotide targets possess a 1 bp 5' overhang which allows for efficient labeling of the target with T4 phage polynucleotide kinase. For the assay the target is added at a concentration of 40 nM or lower (the actual concentration is kept at least 10-fold lower than the lowest protein dilution) and the reaction is allowed to equilibrate for at least 45 min. In addition the

D8 cont

reaction mixture also contains 10 mM Tris (pH 7.5), 100 mM KCl, 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, 5 mM DTT, 10% glycerol, 0.02% BSA (poly (dIdC) or (dAdT) (Pharamacia) can also be added at 10-100 µg/µl).--

---

Please replace the paragraph beginning on page 32, line 29 with the following rewritten paragraph:

D9

--Linker domains between polypeptide domains, e.g., between two ZFPs or between a ZFP and a regulatory domain, can be included. Such linkers are typically polypeptide sequences, such as polyglycine sequences of between about 5 and 200 amino acids. Preferred linkers are typically flexible amino acid subsequences which are synthesized as part of a recombinant fusion protein. For example, in one embodiment, the linker DGGGS (SEQ ID NO:4) is used to link two ZFPs. In another embodiment, the flexible linker linking two ZFPs is an amino acid subsequence comprising the sequence TGEKP (SEQ ID NO:5) (*see, e.g.*, Liu *et al.*, *PNAS* 5525-5530 (1997)). In another embodiment, the linker LRQKDGERP (SEQ ID NO:6) is used to link two ZFPs. In another embodiment, the following linkers are used to link two ZFPs: GGRR (SEQ ID NO:7) (Pomerantz *et al.* 1995, *supra*), $(G_4S)_n$ (SEQ ID NO:8) (Kim *et al.*, *PNAS* 93, 1156-1160 (1996.); and GGRRGGGS (SEQ ID NO:9); LRQRDGERP (SEQ ID NO:10); LRQKDGGGSERP (SEQ ID NO:11); LRQKD$(G_3S)_2$ ERP (SEQ ID NO:12). Alternatively, flexible linkers can be rationally designed using computer program capable of modeling both DNA-binding sites and the peptides themselves (Desjarlais & Berg, *PNAS* 90:2256-2260 (1993), *PNAS* 91:11099-11103 (1994) or by phage display methods.--

---

Please replace the paragraph beginning on page 57, line 8 with the following rewritten paragraph:

D10

--This first Example demonstrates the construction of ZFPs designed to recognize

DNA sequences contained in the promoter of the human vascular endothelial growth factor (VEGF) gene. VEGF is an approximately 46 kDa glycoprotein that is an endothelial cell-specific mitogen induced by hypoxia. VEGF has been implicated in angiogenesis associated with cancer, various retinopathies, and other serious diseases. The DNA target site chosen was a region surrounding the transcription initiation site of the gene. The two 9 base pair (bp) sites chosen are found within the sequence agcGGGGAGGATcGCGGAGGCTtgg (SEQ ID NO:13), where the upper-case letters represent actual 9-bp targets. The protein targeting the upstream 9-bp target was denoted VEGF1, and the protein targeting the downstream 9-bp target was denoted VEGF3a. The major start site of transcription for VEGF is at the T at the 3' end of the first 9-bp target, which is underlined in the sequence above.--

D16 cont

Please replace the paragraph beginning on page 58, line 22 with the following rewritten paragraph:

--VEGF1 (SEQ ID NO:14):

GGTACCCATACCTGGCAAGAAGAAGCAGCACATCTGCCACATCCAGGGCTGT
GGTAAAGTTTACGGCACAACCTCAAATCTGCGTCGTCACCTGCGCTGGCACA
CCGGCGAGAGGCCTTTCATGTGTACCTGGTCCTACTGTGGTAAACGCTTCACC
CGTTCGTCAAACCTGCAGCGTCACAAGCGTACCCACACCGGTGAGAAGAAAT
TTGCTTGCCCGGAGTGTCCGAAGCGCTTCATGCGTAGTGACCACCTGTCCCGT
CACATCAAGACCCACCAGAATAAGAAGGGTGGATCC--

D11

Please replace the paragraph beginning on page 59, line 1 with the following rewritten paragraph:

--VEGF1 translation (SEQ ID NO:15):

VPIPGKKKQHICHIQGCGKVYGTTSNLRRHLRWHTGERPFMCTWSYCGKRFTRS
SNLQRHKRTHTGEKKFACPECPKRFMRSDHLSRHIKTHQNKKGGS--

D12

Please replace the paragraph beginning on page 59, line 4 with the following rewritten paragraph:

--VEGF3a (SEQ ID NO:16):

GGTACCCATACCTGGCAAGAAGAAGCAGCACATCTGCCACATCCAGGGCTGT
GGTAAAGTTTACGGCCAGTCCTCCGACCTGCAGCGTCACCTGCGCTGGCACA
CCGGCGAGAGGCCTTTCATGTGTACCTGGTCCTACTGTGGTAAACGCTTCACC
CGTTCGTCAAACCTACAGAGGCACAAGCGTACACACACCGGTGAGAAGAAAT
TTGCTTGCCCGGAGTGTCCGAAGCGCTTCATGCGAAGTGACGAGCTGTCACG
ACATATCAAGACCCACCAGAACAAGAAGGGTGGATCC--

D 13

Please replace the paragraph beginning on page 59, line 11 with the following rewritten paragraph:

--VEGF3a translation (SEQ ID NO:17):

VPIPGKKKQHICHIQGCGKVYGQSSDLQRHLRWHTGERPFMCTWSYCGKRFTRS
SNLQRHKRTHTGEKKFACPECPKRFMRSDELSRHIKTHQNKKGGS--

D.14

Please replace the paragraph beginning on page 60, line 29 with the following rewritten paragraph:

--VEGF site 1, top: 5'-CATGCATAGCGGGGAGGATCGCCATCGAT (SEQ ID NO:18)

VEGF site 1, bottom: 5'-ATCGATGGCGATCCTCCCCGCTATGCATG (SEQ ID NO:19)

VEGF site 3, top: 5'-CATGCATATCGCGGAGGCTTGGCATCGAT (SEQ ID NO:20)

VEGF site 3, bottom: 5'-ATCGATGCCAAGCCTCCGCGATATGCATG (SEQ ID NO:21)--

D 15

Please replace the paragraph beginning on page 62, line 12 with the following

rewritten paragraph:

--An important consideration in ZFP design is DNA target length. For random DNA, a sequence of n nucleotides would be expected to occur once every 0.5 x 4^n base-pairs. Thus, DNA-binding domains designed to recognize only 9 bp of DNA would find sites every 130,000 bp and could therefore bind to multiple locations in a complex genome (on the order of 20,000 sites in the human genome). 9-bp putative repressor-binding sequences have been chosen for VEGF in the 5' UTR where they might directly interfere with transcription. However, in case zinc finger domains that recognize 9-bp sites lack the necessary affinity or specificity when expressed inside cells, a larger domain was constructed to recognize 18 base-pairs by joining separate three-finger domains with a linker sequence to form a six-finger protein. This should ensure that the repressor specifically targets the appropriate sequence, particularly under conditions where only small amounts of the repressor are being produced. The 9-bp target sites in VEGF were chosen to be adjacent to one another so that the zinc fingers could be linked to recognize an 18-bp sequence. The linker DGGGS (SEQ ID NO:4) was chosen because it permits binding of ZFPs to two 9-bp sites that are separated by a one nucleotide gap, as is the case for the VEGF1 and VEGF3a sites (*see also* Liu *et al., PNAS* 5525-5530 (1997)).--

D16

Please replace the paragraph beginning on page 62, line 28 with the following rewritten paragraph:

--The 6-finger VEGF3a/1 protein encoding sequence was generated as follows. VEGF3a was PCR amplified using the primers SPE7 (5'-GAGCAGAATTCGGCAAGAAGAAGCAGCAC (SEQ ID NO:22)) and SPEamp12 (5'-GTGGTCTAGACAGCTCGTCACTTCGC (SEQ ID NO:23)) to generate EcoRI and XbaI restriction sites at the ends (restriction sites underlined). VEGF1 was PCR amplified using the primers SPEamp13 (5'-GGAGCCAAGGCTGTGGTAAAGTTTACGG (SEQ ID NO:24)) and SPEamp11 (5'-

D17

GGAGAAGCTTGGATCCTCATTATCCC (SEQ ID NO:25)) to generate StyI and HindIII restriction sites at the ends (restriction sites underlined). Using synthetic oligonucleotides, the following sequence was ligated between the XbaI and StyI sites, where XbaI and StyI are underlined: TCT AGA CAC ATC AAA ACC CAC CAG AAC AAG AAA GAC GGC GGT GGC AGC GGC AAA AAG AAA CAG CAC ATA TGT CAC ATC CAA GG (SEQ ID NO:26). This introduced the linker sequence DGGGS (SEQ ID NO:4) between the two SP-1 domains. The ligation product was reamplified with primers SPE7 and SPEamp11 and cloned into pUC19 using the EcoRI and HindIII sites. The linked ZFP sequences were then amplified with primers

(1) GB19

GCCATGCCGGTACCCATACCTGGCAAGAAGAAGCAGCAC (SEQ ID NO:27)

(2) GB10

CAGATCGGATCCACCCTTCTTATTCTGGTGGGT (SEQ ID NO:28) to introduce KpnI and BamHI sites for cloning into the modified pMAL-c2 expression vector as described above.--

Please replace the paragraph beginning on page 63, line 15 with the following rewritten paragraph:

--The nucleotide sequence of the designed, 6-finger ZFP VEGF3a/1 from KpnI to BamHI is:

GGTACCCATACCTGGCAAGAAGAAGCAGCACATCTGCCACATCCAGGGCTGT
GGTAAAGTTTACGGCCAGTCCTCCGACCTGCAGCGTCACCTGCGCTGGCACA
CCGGCGAGAGGCCTTTCATGTGTACCTGGTCCTACTGTGGTAAACGCTTCACA
CGTTCGTCAAACCTACAGAGGCACAAGCGTACACACACAGGTGAGAAGAAA
TTTGCTTGCCCGGAGTGTCCGAAGCGCTTCATGCGAAGTGACGAGCTGTCTAG
ACACATCAAAACCCACCAGAACAAGAAAGACGGCGGTGGCAGCGGCAAAAA
GAAACAGCACATATGTCACATCCAAGGCTGTGGTAAAGTTTACGGCACAACC

TCAAATCTGCGTCGTCACCTGCGCTGGCACACCGGCGAGAGGCCTTTCATGTG
TACCTGGTCCTACTGTGGTAAACGCTTCACCCGTTCGTCAAACCTGCAGCGTC
ACAAGCGTACCCACACCGGTGAGAAGAAATTTGCTTGCCCGGAGTGTCCGAA
GCGCTTCATGCGTAGTGACCACCTGTCCCGTCACATCAAGACCCACCAGAAT
AAGAAGGGTGGATCC (SEQ ID NO:29)--

Please replace the paragraph beginning on page 63, line 29 with the following rewritten paragraph:

--The VEGF3a/1 amino acid translation (using single letter code) is:
VPIPGKKKQHICHIQGCGKVYGQSSDLQRHLRWHTGERPFMCTWSYCGKRFTRS
SNLQRHKRTHTGEKKFACPECPKRFMRSDELSRHIKTHQNKKDGGGSGKKKQHI
CHIQGCGKVYGTTSNLRRHLRWHTGERPFMCTWSYCGKRFTRSSNLQRHKRTH
TGEKKFACPECPKRFMRSDHLSRHIKTHQNKKGGS (SEQ ID NO:30)--

Please replace the paragraph beginning on page 64, line 1 with the following rewritten paragraph:

--The 18-bp binding protein VEGF3a/1 was expressed in *E. coli* as an MBP fusion, purified by affinity chromatography, and tested in EMSA experiments as described in Example 1. The target oligonucleotides were prepared as described and comprised the following complementary sequences:

(1) JVF9
AGCGAGCGGGGAGGATCGCGGAGGCTTGGGGCAGCCGGGTAG (SEQ ID NO:31), and

(2) JVF10
CGCTCTACCCGGCTGCCCCAAGCCTCCGCGATCCTCCCCGCT (SEQ ID NO:32).--

Please replace the paragraph beginning on page 65, line 23 with the following rewritten paragraph:

--The VP16 protein of HSV-1 has been studied extensively, and it has been shown that the C-terminal 78 amino acids can act as a trans-activation domain when fused to a DNA-binding domain (Hagmann *et al., J. Virology* 71:5952-5962 (1997)). VP16 has also been shown to function at a distance and in an orientation-independent manner. For these studies, amino acids 413 to 490 in the VP16 protein sequence were used. DNA encoding this domain was PCR amplified from plasmid pMSVP16ΔC+119 using primers with the following sequences:

(1) JVF24

CGCGGATCCGCCCCCCCGACCGATG (SEQ ID NO:33), and

(2) JVF25

CCGCAAGCTTACTTGTCATCGTCGTCCTTGTAGTCGCTGCCCCCACCGTACTC GTCAATTCC (SEQ ID NO:34).--

Please replace the paragraph beginning on page 66, line 5 with the following rewritten paragraph:

--Three expression vectors were constructed for these studies. The general design is summarized in Figure 5. The vectors are derived from pcDNA3.1(+) (Invitrogen), and place the ZFP constructs under the control of the cytomegalovirus (CMV) promoter. The vector carries ampicillin and neomycin markers for selection in bacteria and mammalian cell culture, respectively. A Kozak sequence for proper translation initiation (Kozak, *J. Biol. Chem.* 266:19867-19870 (1991)) was incorporated. To achieve nuclear localization of the products, the nuclear localization sequence (NLS) from the SV40 large T antigen (Pro-Lys-Lys-Lys-Arg-Lys-Val (SEQ ID NO:35)) (Kalderon *et al., Cell* 39:499-509 (1984)) was added. The insertion site for the ZFP-encoding sequence is followed by the functional domain sequence. The three versions of this vector differ in the functional

D22 cont

domain; "pcDNA-NKF" carries the KRAB repression domain sequence, "pcDNA-NVF" carries the VP16 activation domain, and "NF-control" carries no functional domain. Following the functional domain is the FLAG epitope sequence (Kodak) to allow specific detection of the ZFPs.--

Please replace the paragraph beginning on page 66, line 19 with the following rewritten paragraph:

D23

--The vectors were constructed as follows. Plasmid pcDNA-ΔHB was constructed by digesting plasmid pcDNA3.1(+) (Invitrogen) with HindIII and BamHI, filling in the sticky ends with Klenow, and religating. This eliminated the HindIII, KpnI, and BamHI sites in the polylinker. The vector pcDNA3.1(+) is described in the Invitrogen catalog. Plasmid pcDNA-NKF was generated by inserting a fragment into the EcoRI/XhoI sites of pcDNA-ΔHB that contained the following: 1) a segment from EcoRI to KpnI containing the Kozak sequence including the initiation codon and the SV40 NLS sequence, altogether comprising the DNA sequence

GAATTCGCTAGCGCCACCATGGCCCCCAAGAAGAAGAGGAAGGTGGGAATC CATGGGGTAC (SEQ ID NO:36),

where the EcoRI and KpnI sites are underlined; and 2) a segment from KpnI to XhoI containing a BamHI site, the KRAB-A box from KOX1 (amino acid coordinates 11-53 in Thiesen, 1990, *supra*), the FLAG epitope (from Kodak/IBI catalog), and a HindIII site, altogether comprising the sequence

GGTACCCGGGGATCCCGGACACTGGTGACCTTCAAGGATGTATTTGTGGACT TCACCAGGGAGGAGTGGAAGCTGCTGGACACTGCTCAGCAGATCGTGTACAG AAATGTGATGCTGGAGAACTATAAGAACCTGGTTTCCTTGGGCAGCGACTAC AAGGACGACGATGACAAGTAAGCTTCTCGAG (SEQ ID NO:37)

where the KpnI, BamHI and XhoI sites are underlined.--

Please replace the paragraph beginning on page 67, line 14 with the following rewritten paragraph:

--The effector plasmids used in Example V were constructed as follows. Plasmid pcDNA-NVF was constructed by PCR amplifying the VP16 transactivation domain, as described above, and inserting the product into the BamHI/HindIII sites of pcDNA-NKF, replacing the KRAB sequence. The sequence of the inserted fragment, from BamHI to HindIII, was:

D24

GGATCCGCCCCCCCGACCGATGTCAGCCTGGGGGACGAGCTCCACTTAGACG
GCGAGGACGTGGCGATGGCGCATGCCGACGCGCTAGACGATTTCGATCTGGA
CATGTTGGGGGACGGGGATTCCCCGGGGCCGGGATTTACCCCCCACGACTCC
GCCCCCTACGGCGCTCTGGATATGGCCGACTTCGAGTTTGAGCAGATGTTTAC
CGATGCCCTTGGAATTGACGAGTACGGTGGGGGCAGCGACTACAAGGACGAC
GATGACAAGTAAGCTT (SEQ ID NO:38).--

Please replace the paragraph beginning on page 67, line 28 with the following rewritten paragraph:

--The effector plasmids used in Example VI were constructed as follows. Plasmid NF-control was generated by inserting the sequence

D25

GAATTCGCTAGCGCCACCATGGCCCCCAAGAAGAAGAGGAAGGTGGGAATC
CATGGGGTACCCGGGGATGGATCCGGCAGCGACTACAAGGACGACGATGAC
AAGTAAGCTTCTCGAG (SEQ ID NO:39)

into the EcoRI-XhoI sites of pcDNA-NKF, thereby replacing the NLS-KRAB-FLAG sequences with NLS-FLAG only.--

Please replace the paragraph beginning on page 68, line 31 with the following rewritten paragraph:

D26

--The reporter plasmid system was based on the pGL3 firefly luciferase vectors